Compressed Sensing Tools for Radial Velocity Data Analysis

Nathan Hara

Observatoire de Paris, IMCCE

With Jacques Laskar, Gwenaël Boué, Alexandre Correia

University of Warwick, September 7$^{\text{th}}$ 2017

CELMECVII, San Martino al Cimino (Viterbo, Italy), 3-9 september 2017

AMD-stability and the classification of planetary systems

Jacques Laskar

with Antoine Petit Gwenaël Boué

IMCCE, Observatoire de Paris

Laskar, On the spacing of planetary systems, 2000, PRL

Laskar, Petit, AMD-stability and the classification of planetary systems, 2017, A&A

Petit, Laskar, Boué, AMD-stability in the presence of first-order mean motion resonances, 2017, A&A

Overview

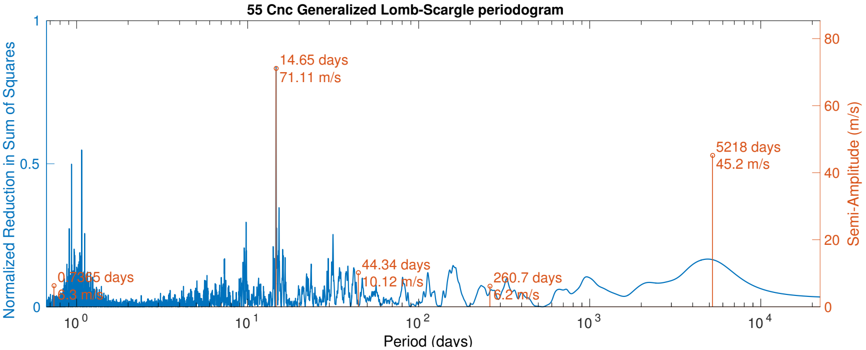
1) Angular Momentum Deficit

2) Avoiding spurious periodogram peaks with ℓ_1 minimization.

3) Comments on the radial velocity challenge (Dumusque 2016, Dumusque et al 2017)

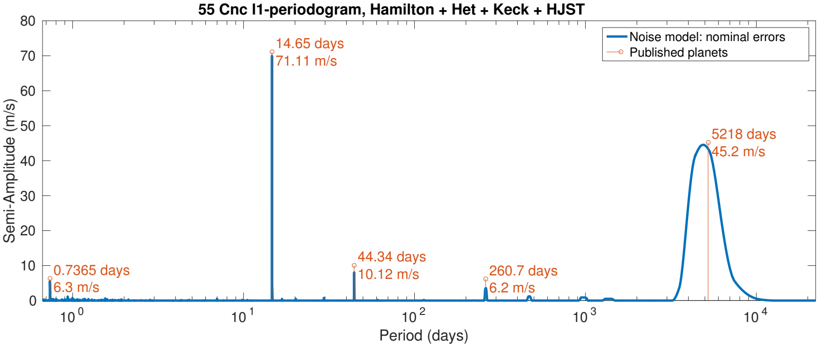
Hara, Boué, Laskar, Correia 2017

Motivation



55 Cnc Generalized Lomb-Scargle periodogram
Normalized Reduction in Sum of Squares
Semi-Amplitude (m/s)
Period (days)
14.65 days
71.11 m/s
5218 days
45.2 m/s
0.7365 days
6.3 m/s
44.34 days
10.12 m/s
260.7 days
6.2 m/s

Motivation



55 Cnc l1-periodogram, Hamilton + Het + Keck + HJST
Noise model: nominal errors
Published planets
14.65 days
71.11 m/s
5218 days
45.2 m/s
0.7365 days
6.3 m/s
44.34 days
10.12 m/s
260.7 days
6.2 m/s
Semi-Amplitude (m/s)
Period (days)

Advertising

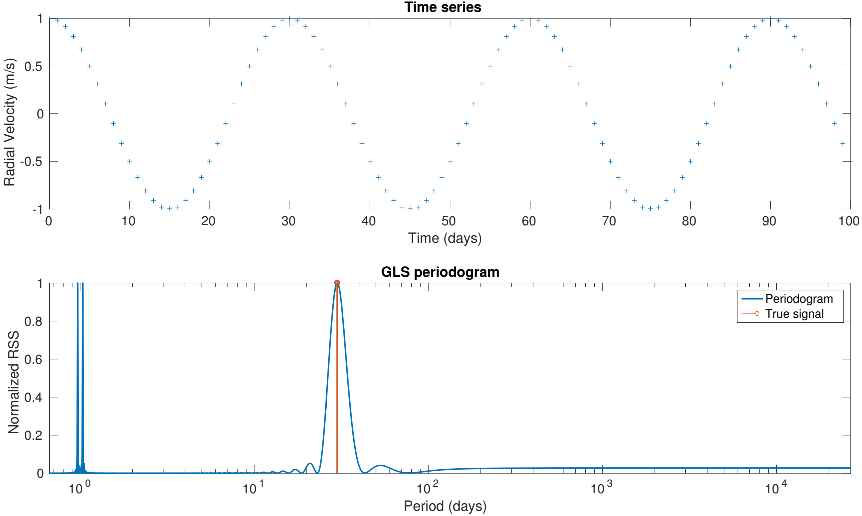
Periodogram : Lomb 1976, Ferraz-Mello 1981, Scargle 1982, Zechmeister et Kürster 2009

$$A_\omega, B_\omega, C_\omega = \min_{A,B,C} \|y(t) - A\cos(\omega t) - B\sin(\omega t) - C\| \tag{1}$$

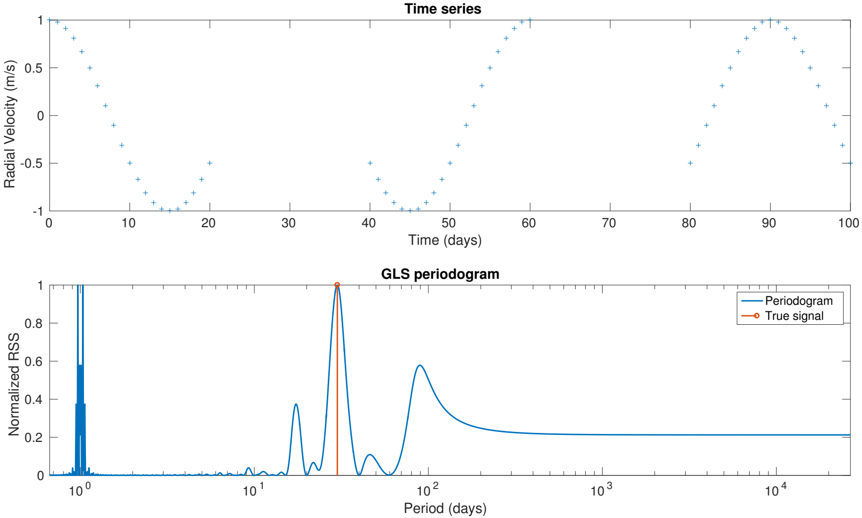
$$P(\omega) = \frac{\|y(t)\|^2 - \|y(t) - A_\omega \cos(\omega t) - B_\omega \sin(\omega t) - C_\omega\|^2}{\|y(t)\|^2} \tag{2}$$

For all ω, there is an explicit solution to this linear least square problem.

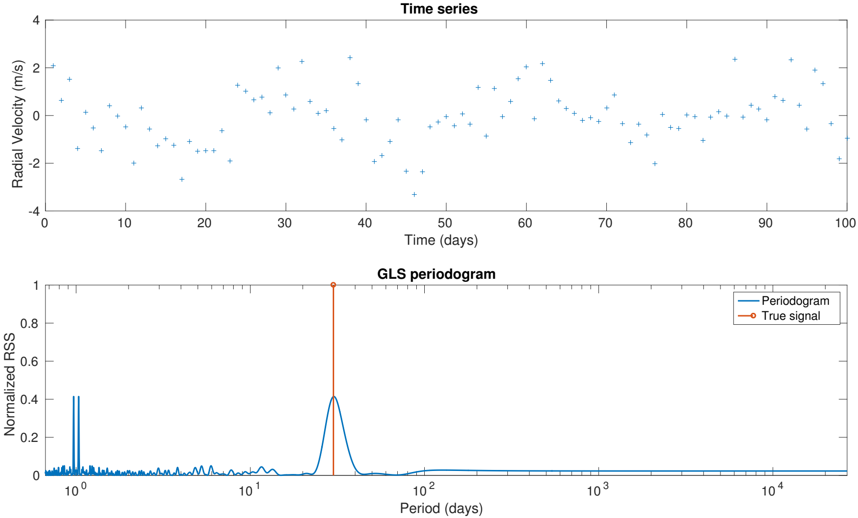
Exemple : regular sampling, no noise



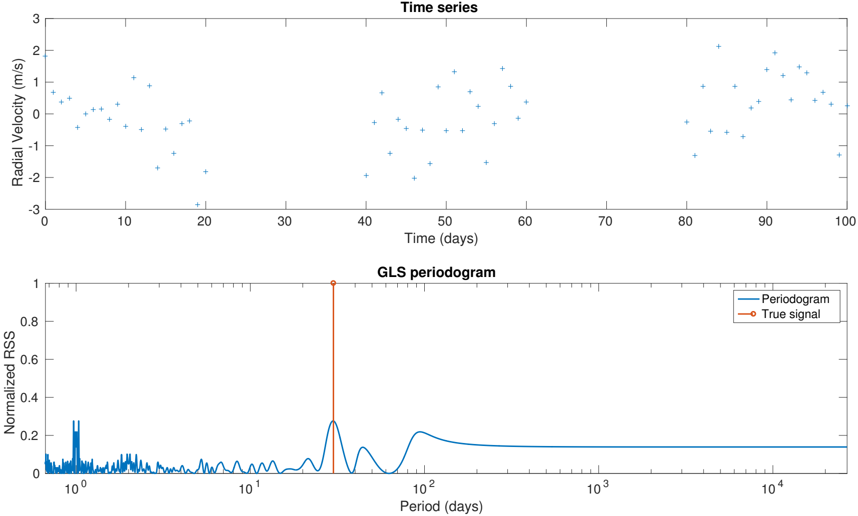
Exemple : missing samples



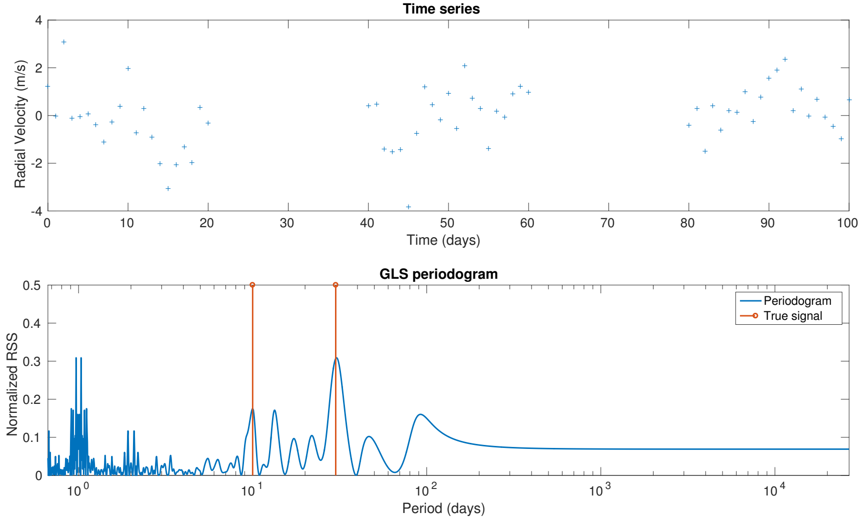
Exemple : regular sampling plus noise



Exemple : missing samples plus noise

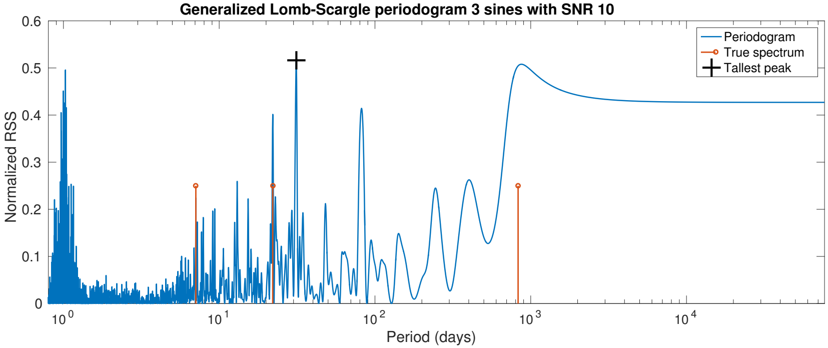


Exemple : missing samples plus noise, two signals



The periodogram can be wrong

$$P(\omega) = \frac{\|y(t)\|^2 - \|y(t) - A_{\text{fit}}\cos(\omega t + \phi_{\text{fit}}) - C_{\text{fit}}\|^2}{\|y(t)\|^2} \qquad (3)$$



The periodogram looks at one frequency at a time : can we search for several at once? and avoid lengthy random searches?

Workarounds

Search for several planets at a time

- Bayesian analysis with Fusion MCMC (Campbell & Walker 1998, Gregory 2005), parallel tempering (Swendsen & Wang 1986), Nested sampling (Brouwer 2005, Faria 2014...).
- Genetic algorithm (Ségransan 2011).
- Multi-frequency periodogram (Baluev 2013).
- MMSE i.e. brute force discretization (Jenkins 2014).

Workarounds

Search for several planets at a time

- Bayesian analysis with Fusion MCMC (Campbell & Walker 1998, Gregory 2005), parallel tempering (Swendsen & Wang 1986), Nested sampling (Brouwer 2005, Faria 2014...).
- Genetic algorithm (Ségransan 2011).
- Multi-frequency periodogram (Baluev 2013).
- MMSE i.e. brute force discretization (Jenkins 2014).

Or consider the whole Fourier spectrum as the unknown.

An overcomplete representation

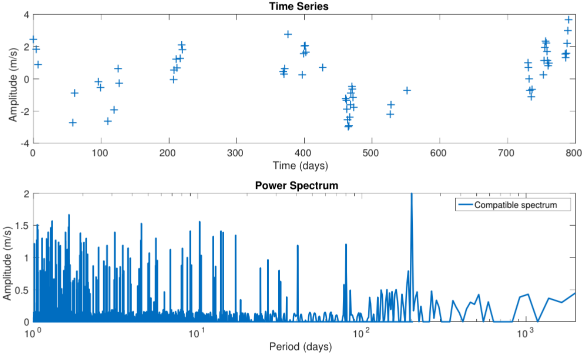
Linear equation

- $\boldsymbol{y}(\boldsymbol{t})$: vector of m observations
- $\mathbf{A}$: matrix whose columns are sine functions
- $\boldsymbol{x}$: Fourier spectrum
- ϵ noise

$$\boldsymbol{y}(\boldsymbol{t}) = \mathbf{A}\boldsymbol{x} + \boldsymbol{\epsilon} \tag{4}$$

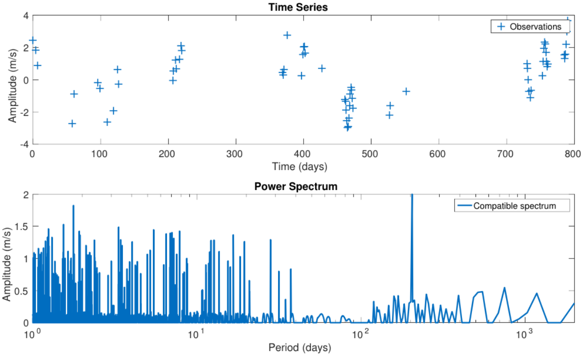
$$\begin{pmatrix} y(t_1) \\ \vdots \\ y(t_m) \end{pmatrix} = \begin{pmatrix} e^{-\omega_1 t_1} & \dots & e^{-\omega_n t_1} & e^{\omega_1 t_1} & \dots & e^{\omega_n t_1} \\ \vdots & & & & & \vdots \\ e^{-\omega_1 t_m} & \dots & e^{-\omega_n t_m} & e^{\omega_1 t_m} & \dots & e^{\omega_n t_m} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ x_{n+1} \\ \vdots \\ x_{2n} \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_m \end{pmatrix} \tag{5}$$

An overcomplete representation



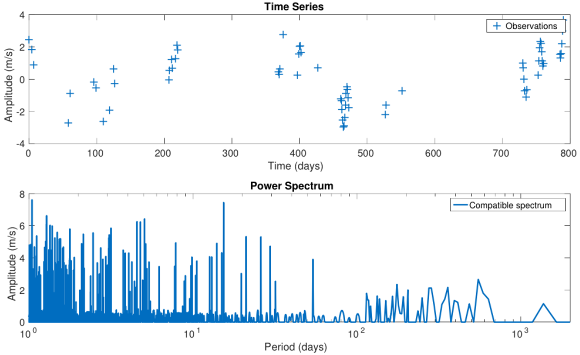
$$\begin{pmatrix} y(t_1) \\ \vdots \\ y(t_m) \end{pmatrix} = \begin{pmatrix} e^{-\omega_1 t_1} & \dots & e^{-\omega_n t_1} & e^{\omega_1 t_1} & \dots & e^{\omega_n t_1} \\ \vdots & & & & & \vdots \\ e^{-\omega_1 t_m} & \dots & e^{-\omega_n t_m} & e^{\omega_1 t_m} & \dots & e^{\omega_n t_m} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ x_{n+1} \\ \vdots \\ x_{2n} \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_m \end{pmatrix} \tag{4}$$

An overcomplete representation



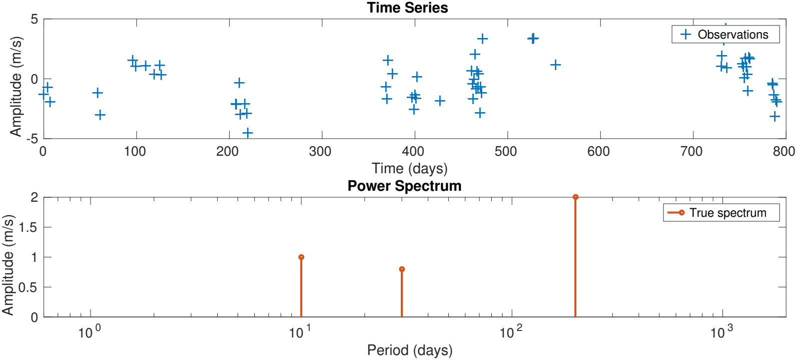
$$\begin{pmatrix} y(t_1) \\ \vdots \\ y(t_m) \end{pmatrix} = \begin{pmatrix} e^{-\omega_1 t_1} & \dots & e^{-\omega_n t_1} & e^{\omega_1 t_1} & \dots & e^{\omega_n t_1} \\ \vdots & & & & & \vdots \\ e^{-\omega_1 t_m} & \dots & e^{-\omega_n t_m} & e^{\omega_1 t_m} & \dots & e^{\omega_n t_m} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ x_{n+1} \\ \vdots \\ x_{2n} \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_m \end{pmatrix} \tag{4}$$

An overcomplete representation



$$\begin{pmatrix} y(t_1) \\ \vdots \\ y(t_m) \end{pmatrix} = \begin{pmatrix} e^{-\omega_1 t_1} & \dots & e^{-\omega_n t_1} & e^{\omega_1 t_1} & \dots & e^{\omega_n t_1} \\ \vdots & & & & & \vdots \\ e^{-\omega_1 t_m} & \dots & e^{-\omega_n t_m} & e^{\omega_1 t_m} & \dots & e^{\omega_n t_m} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ x_{n+1} \\ \vdots \\ x_{2n} \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_m \end{pmatrix} \tag{4}$$

An overcomplete representation



$$\begin{pmatrix} y(t_1) \\ \vdots \\ y(t_m) \end{pmatrix} = \begin{pmatrix} e^{-\omega_1 t_1} & \dots & e^{-\omega_n t_1} & e^{\omega_1 t_1} & \dots & e^{\omega_n t_1} \\ \vdots & & & & & \vdots \\ e^{-\omega_1 t_m} & \dots & e^{-\omega_n t_m} & e^{\omega_1 t_m} & \dots & e^{\omega_n t_m} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ x_k \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_m \end{pmatrix} \quad (4)$$

An overcomplete representation

- $\boldsymbol{y}(\boldsymbol{t})$: vector of m observations
- $\mathbf{A}$: matrix whose columns are sine functions
- x : true spectrum of $y(t)$
- ϵ noise

$$\boldsymbol{y}(\boldsymbol{t}) = \mathbf{A}\boldsymbol{x} + \epsilon \tag{4}$$

A few entries of $\boldsymbol{x}$ are non-zero

$$\begin{pmatrix} y(t_1) \\ \vdots \\ y(t_m) \end{pmatrix} = \begin{pmatrix} e^{-\omega_1 t_1} & \dots & e^{-\omega_n t_1} & e^{\omega_1 t_1} & \dots & e^{\omega_n t_1} \\ \vdots & & & & & \vdots \\ e^{-\omega_1 t_m} & \dots & e^{-\omega_n t_m} & e^{\omega_1 t_m} & \dots & e^{\omega_n t_m} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ x_k \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_m \end{pmatrix} \tag{5}$$

Sparse representations

$$\boldsymbol{y}(\boldsymbol{t}) = \mathbf{A}\boldsymbol{x} + \epsilon, \;\; \boldsymbol{y} \in \mathbb{R}^m, \; \boldsymbol{x} \in \mathbb{R}^n, \; n >> m, \;\; \boldsymbol{x} \text{ has many zeros}$$

What is $\boldsymbol{x}$?

First idea

Trying to find the $\boldsymbol{x}$ that reproduces the data $\mathbf{A}\boldsymbol{x} = \boldsymbol{y}$ that has the maximum number of coefficients equal to zero.

Problems

- That is a NP-hard problem (Ge et al 2011)
- If there is noise, why trying to reproduce exactly the data ?

Sparse representations

$$\boldsymbol{y}(\boldsymbol{t}) = \mathbf{A}\boldsymbol{x} + \epsilon, \ \ \boldsymbol{y} \in \mathbb{R}^m, \ \boldsymbol{x} \in \mathbb{R}^n, \ n >> m, \ \ \boldsymbol{x} \text{ has many zeros}$$

What is $\boldsymbol{x}$?

Astronomy	Compressed Sensing
CLEAN (Roberts et al 1987)	Matching Pursuit (Mallat & Zhang 1993)
Frequency Map Analysis (Laskar 1988, 1990), CLEANest (Foster 1955)	Orthogonal matching pursuit (Tropp et al 2007)
IAA (Babu & Stoica 2010)	Iteratively re-weighted least square (Daubechies et al 2010)
Bourguignon et al 2007, Hara et al 2017	Basis pursuit (Chen and Donoho 1993), Lasso (Tibshirani et al 1994)

Basis pursuit algorithms (Chen & Donoho 1998)

$$\boldsymbol{x}^\star = \underset{\boldsymbol{x}\in\mathbb{C}^n}{\arg\min} \quad \|\boldsymbol{x}\|_{\ell_1} \quad \text{s. t.} \quad \|\mathbf{W}(\mathbf{A}\boldsymbol{x} - \boldsymbol{y}(\boldsymbol{t}))\|_{\ell_2} \leqslant \epsilon \qquad (\mathrm{BP}_{\epsilon,W})$$

- $\|\boldsymbol{x}\|_{\ell_1} = \sum_{k=1}^{2n} |x_k|$
- $\boldsymbol{y}(\boldsymbol{t})$: vector of m observations
- $\mathbf{A} = (\mathrm{e}^{\mathrm{i}\omega t})_{\omega=\omega_1..\omega_n}$: matrix whose columns are sine functions
- ϵ : tolerance, sets the **trade-off between the observations and the sparsity of the Fourier spectrum**

Basis pursuit algorithms (Chen & Donoho 1998)

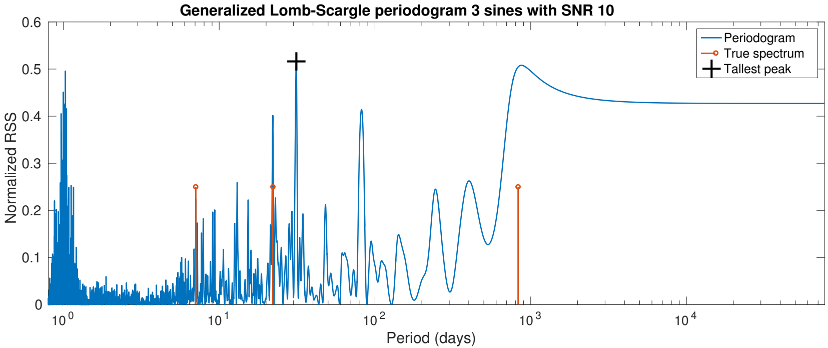
$$\boldsymbol{x}^\star = \arg\min_{\boldsymbol{x}\in\mathbb{C}^n} \quad \|\boldsymbol{x}\|_{\ell_1} \quad \text{s. t.} \quad \|\mathbf{W}(\mathbf{A}\boldsymbol{x} - \boldsymbol{y}(\boldsymbol{t}))\|_{\ell_2} \leqslant \epsilon \qquad (\mathrm{BP}_{\epsilon,W})$$

- $\|\boldsymbol{x}\|_{\ell_1} = \sum_{k=1}^{2n} |x_k|$
- $\boldsymbol{y}(\boldsymbol{t})$: vector of m observations
- $\mathbf{A} = (\mathrm{e}^{\mathrm{i}\omega t})_{\omega=\omega_1..\omega_n}$: matrix whose columns are sine functions
- ϵ : tolerance, sets the trade-off between the observations and the sparsity of the Fourier spectrum

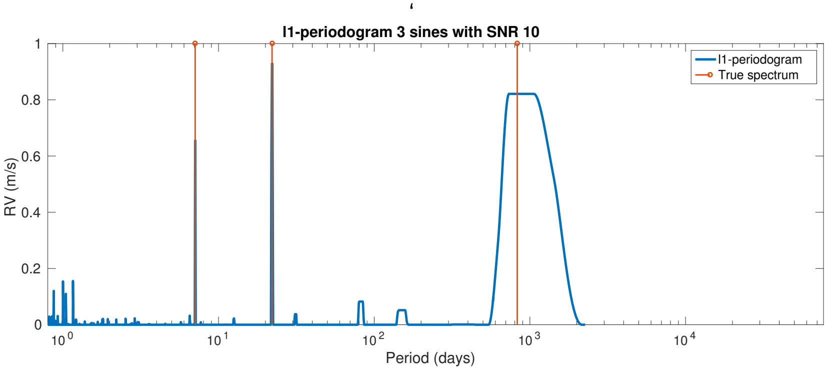
- Automatic tuning of the algorithm parameters
- $\mathbf{W}$: weight matrix $m \times m$ (for correlated Gaussian noise/Gaussian process)
- Plus scaling and smoothing for atomic norm de-noising
- Assessment of the peak statistical significance

Convex problem : fast algorithms

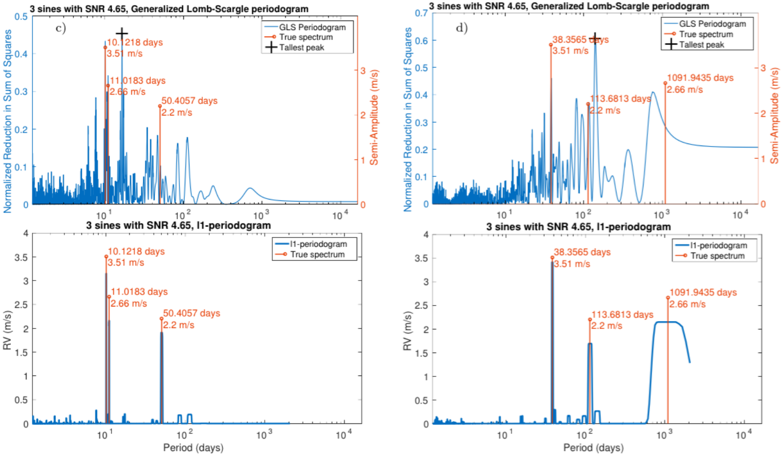
Back to the example



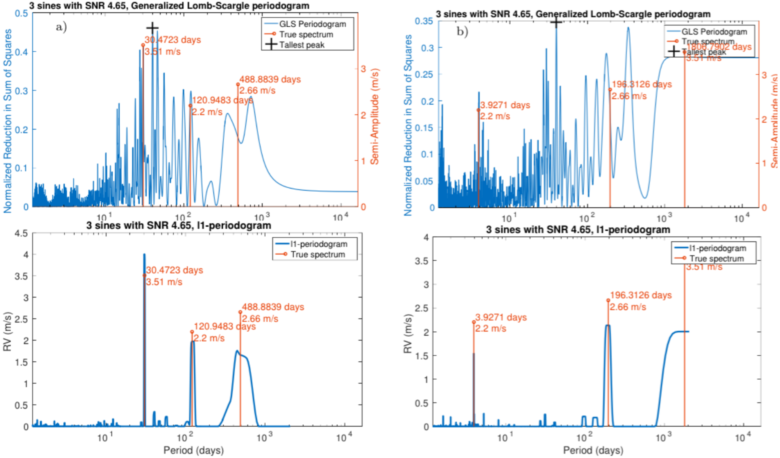
Back to the example



Examples of first wrong peak

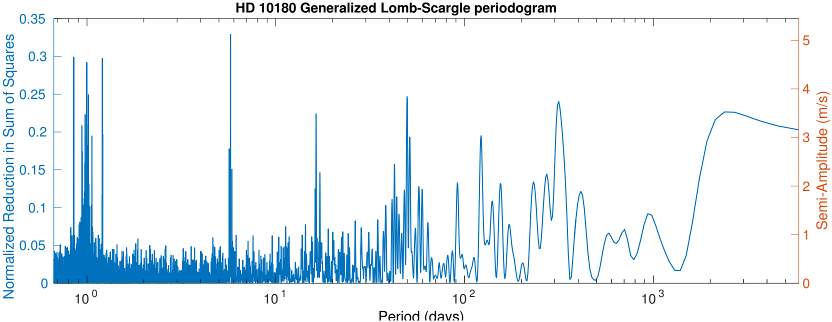


Examples of first wrong peak



HD 10180

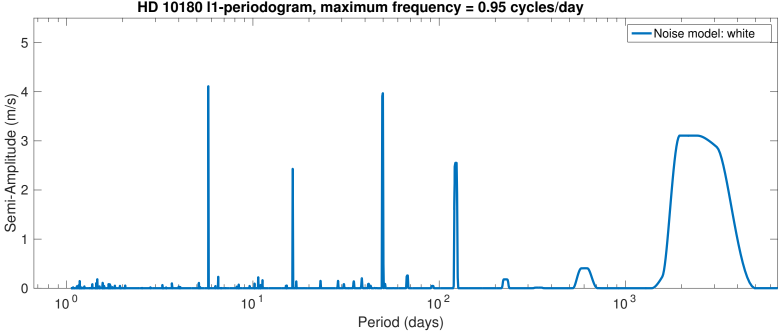
190 observations
Precision ≈ 1 m.s^{-1}



HD 10180 (Lovis et al 2011)

190 observations

Precision $\approx$ 1 m.s^{-1}

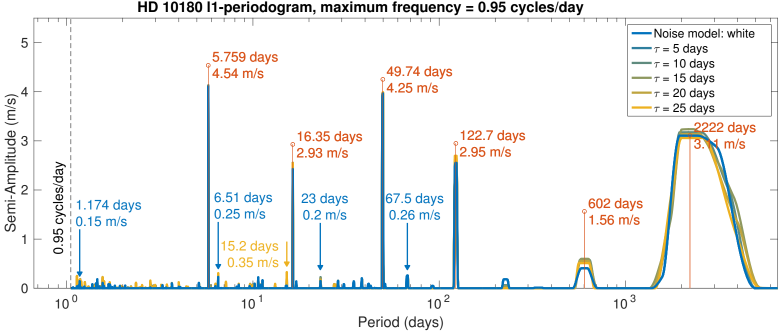


Hara, Boué, Laskar, Correia 2011

HD 10180 ℓ_1-periodogram

190 observations

Précision $\approx$ 1 m.s^{-1}

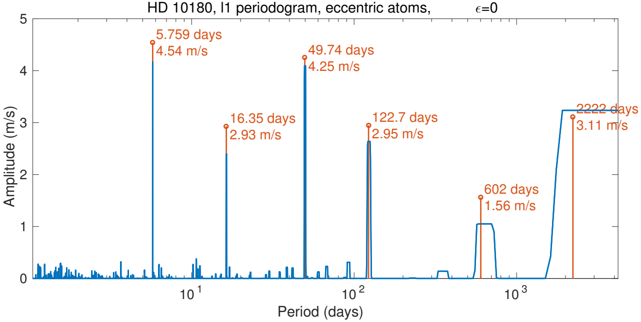


Hara, Boué, Laskar, Correia 2017

HD 10180 ℓ_1-periodogram with Keplerian representation

190 observations

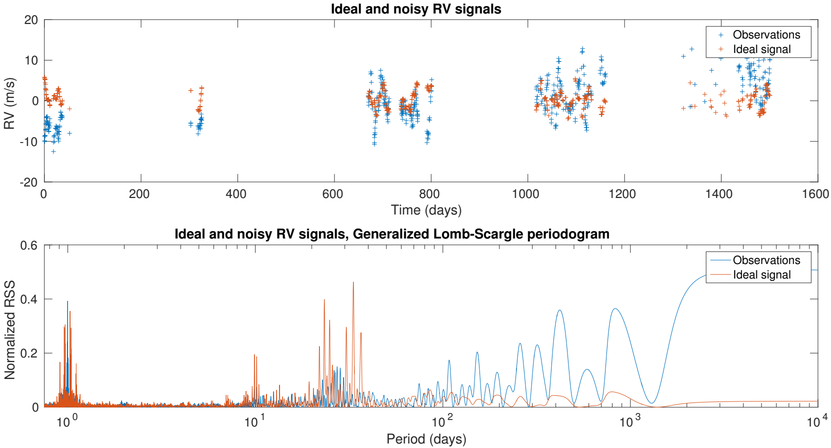
Precision $\approx$ 1 m.s^{-1}

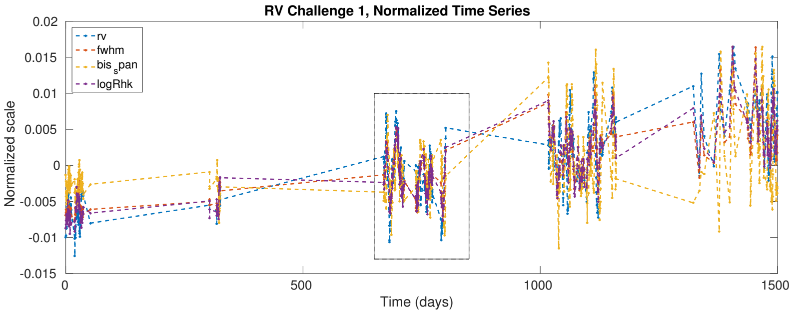


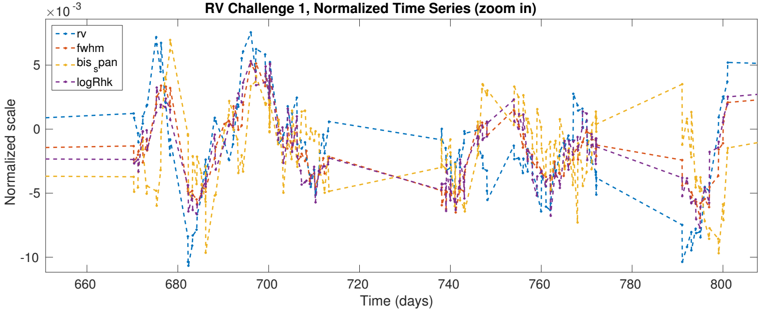
Radial Velocity Fitting Challenge

- 15 systems generated with SOAP 2.0 (Dumusque et al 2014), **dominated by activity**
- Between 100 and 500 high precision measurements
- 8 teams analysed the results
- See Dumusque (2016) for details

Activity dominated (blue) and ideal signal, no noises (red)



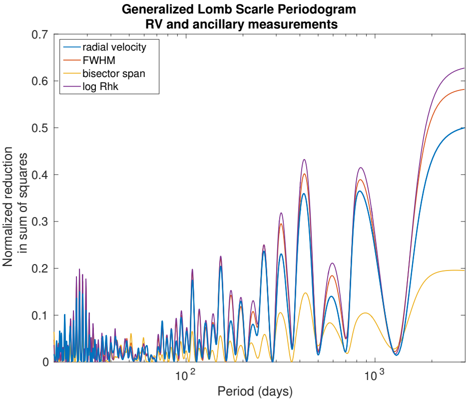

RV Challenge 1, Normalized Time Series
Normalized scale
Time (days)
rv
fwhm
bis span
logRhk
0.02
0.015
0.01
0.005
0
-0.005
-0.01
-0.015
0
500
1000
1500

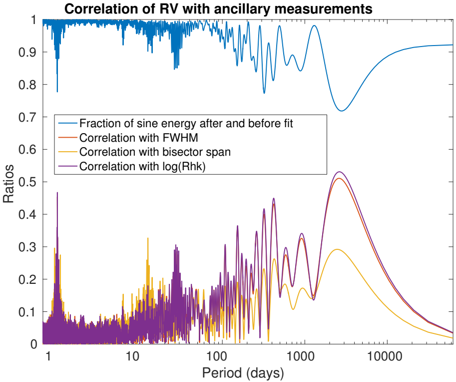

RV Challenge 1, Normalized Time Series (zoom in)
×10 -3
Normalized scale
Time (days)
rv
fwhm
bis span
logRhk
5
0
-5
-10
660
680
700
720
740
760
780
800

Data analysis method

Fitting a linear combination of FWHM, Bisector span and $\log R'_{HK}$ to the data plus a constant and a trend.
Similar in spirit to Queloz 2001, Boisse 2009, Robertson (2014, 2015)

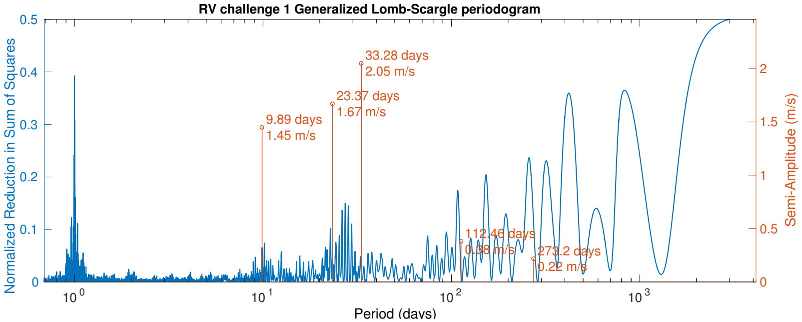
Performing Lomb-Scargle periodogram and ℓ_1 periodogram on the residuals.



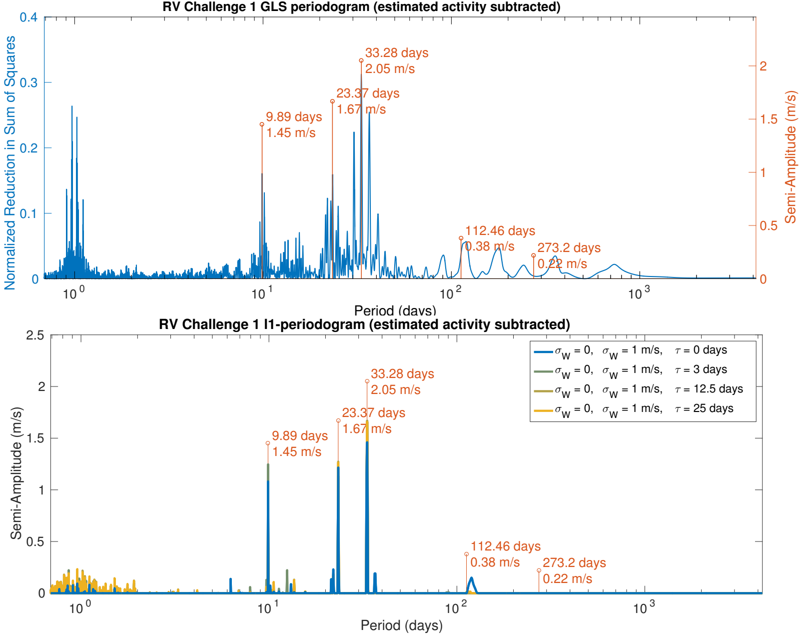


$$\text{Fraction of the sine absorbed}(\omega) = \frac{(y_\omega - y_{fit})^T V^{-1} (y_\omega - y_{fit})}{y_\omega^T V^{-1} y_\omega} \qquad (6)$$

RV fitting challenge, system 1 : Generalised Lomb-Scargle periodogram of raw data

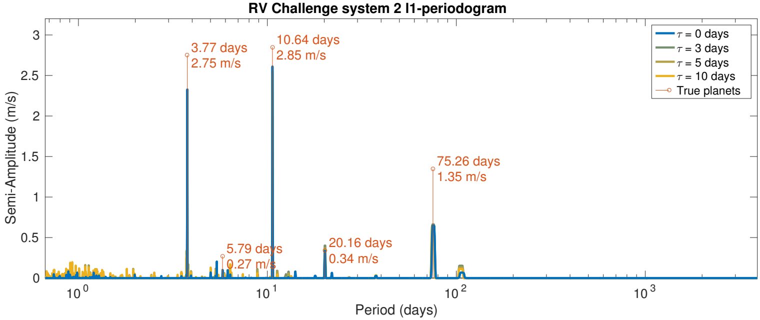


Hara, Boué, Laskar, Correia 2016

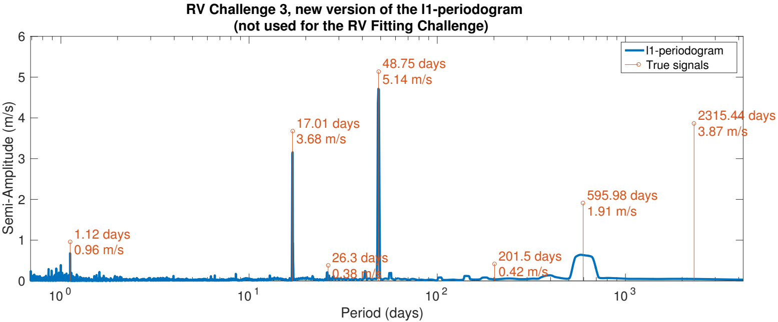


Hara, Boué, Laskar, Correia 2016

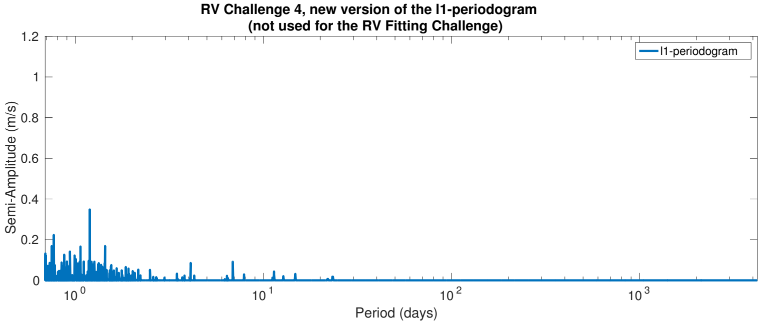
Radial velocity challenge system 2



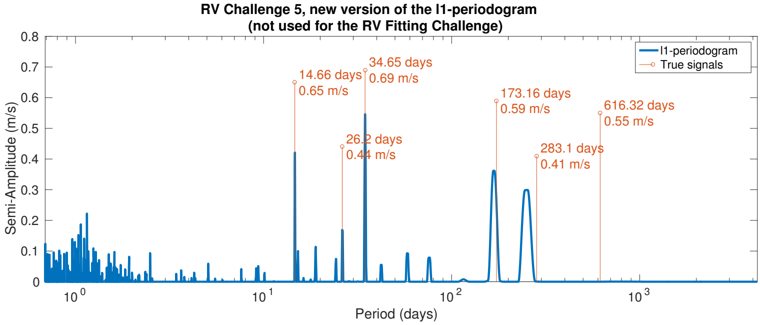
Radial velocity challenge system 3



Radial velocity challenge system 4



Radial velocity challenge system 5



Conclusion

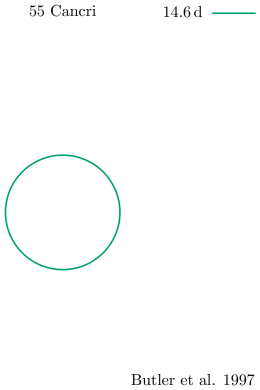
- ℓ_1 minimization is an efficient way to search for several planets at once. Drastically reduces the number of spurious peaks on the periodogram.
- Not as precise as an efficient brute force exploration but much faster ($\approx 30s$ for 200 measurement points).

- Estimating the activity as a linear combination of FWHM, BIS and $\log R'_{HK}$, a constant and a trend works extremely well on the radial velocity fitting challenge.
- Is it true on real measurements ?

Hara, Boué, Laskar & Correia 2017

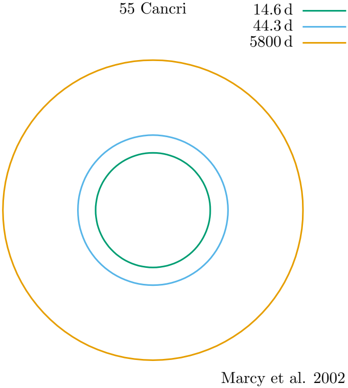
Appendices

55 Cancri



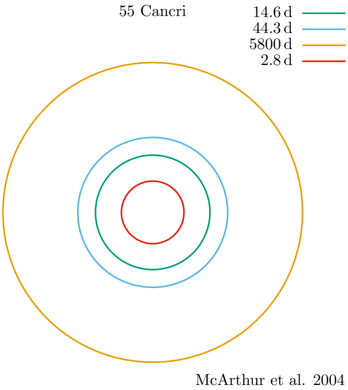
Planet at 14.6 days minimum mass of 0.8 M_j (Butler, Marcy, Williams, Hauser, Shirts 1997)

55 Cancri



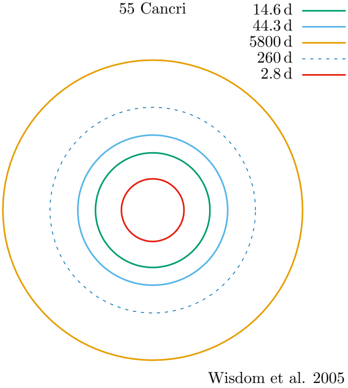
Planet ≈ 5800 days and a Jupiter at 44.3 days (Marcy, Butler, Fischer, Laughlin, Vogt, Henry, Pourbaix 2002)

55 Cancri



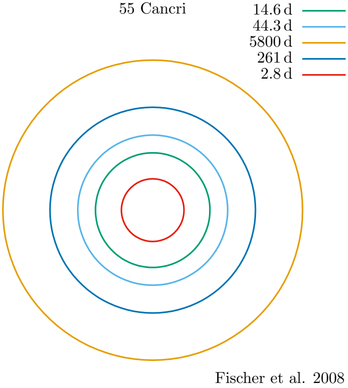
McArthur et al (2004) suggests there might be a a Neptune with an orbital period of 2.8 days

55 Cancri



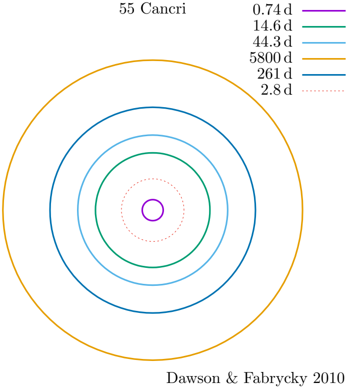
Wisdom (2005) analyses the same data set and suggests a planet at 260 days

55 Cancri



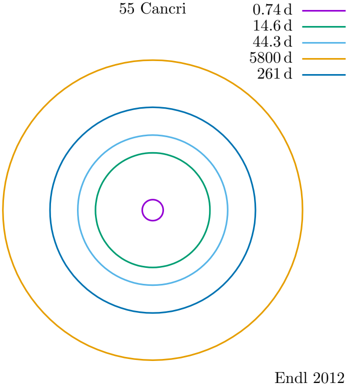
Fischer et al (2008) confirms a 261 days Neptune with additional measurements

55 Cancri



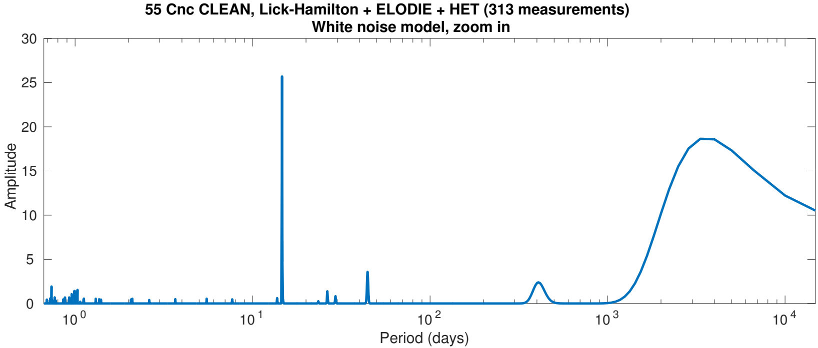
Dawson & Fabricky (2010) shows the 2.8 days periodicity is an alias of a 0.7365 days period (whose transit was observed, Winn et al 2011)

55 Cancri



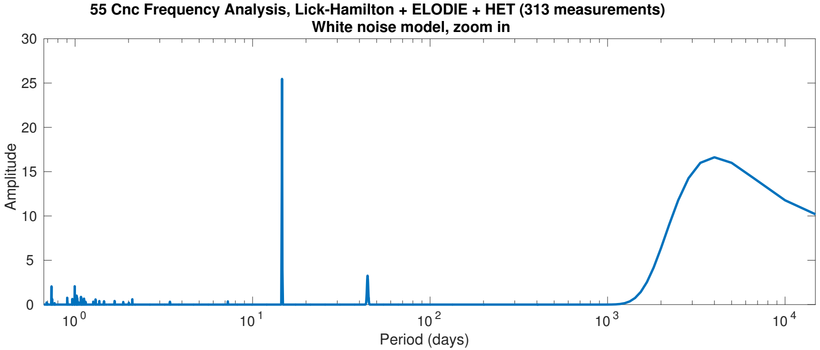
Endl, Robertson, Cochran, MacQueen (2012) refines the orbital parameters estimates

55 Cancri



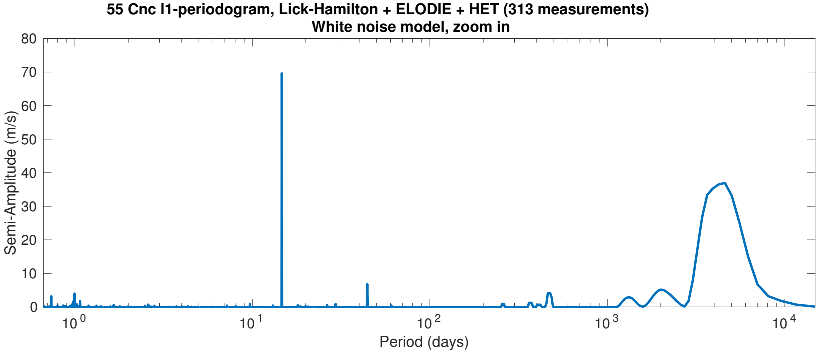
55 Cnc CLEAN, Lick-Hamilton + ELODIE + HET (313 measurements)
White noise model, zoom in
Amplitude
Period (days)
30
25
20
15
10
5
0
10^0
10^1
10^2
10^3
10^4

55 Cancri



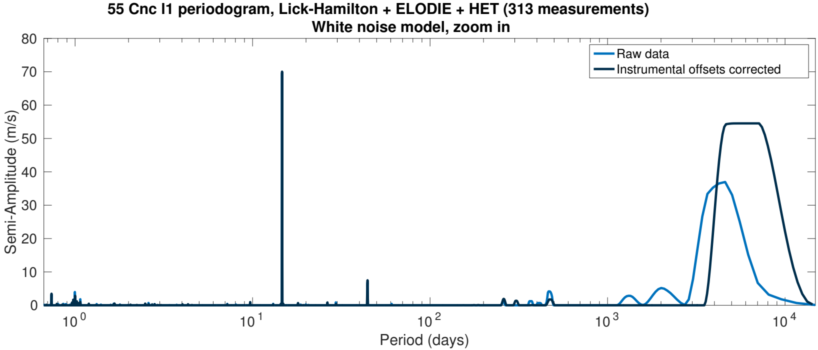
55 Cnc Frequency Analysis, Lick-Hamilton + ELODIE + HET (313 measurements)
White noise model, zoom in
Amplitude
30
25
20
15
10
5
0
10^0
10^1
10^2
10^3
10^4
Period (days)

55 Cancri



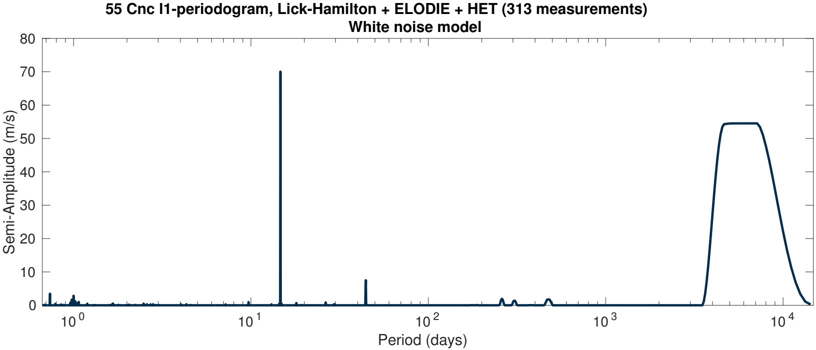
55 Cnc l1-periodogram, Lick-Hamilton + ELODIE + HET (313 measurements)
White noise model, zoom in
Semi-Amplitude (m/s)
0
10
20
30
40
50
60
70
80
10^0
10^1
10^2
10^3
10^4
Period (days)

55 Cancri



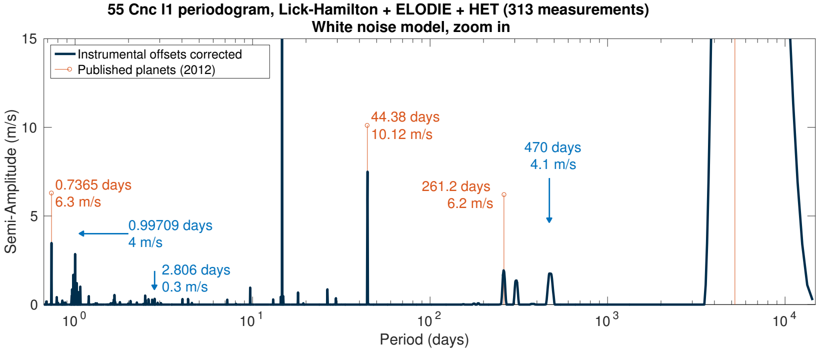
55 Cnc I1 periodogram, Lick-Hamilton + ELODIE + HET (313 measurements)
White noise model, zoom in
Raw data
Instrumental offsets corrected
Semi-Amplitude (m/s)
Period (days)
0
10
20
30
40
50
60
70
80
10^0
10^1
10^2
10^3
10^4

55 Cancri

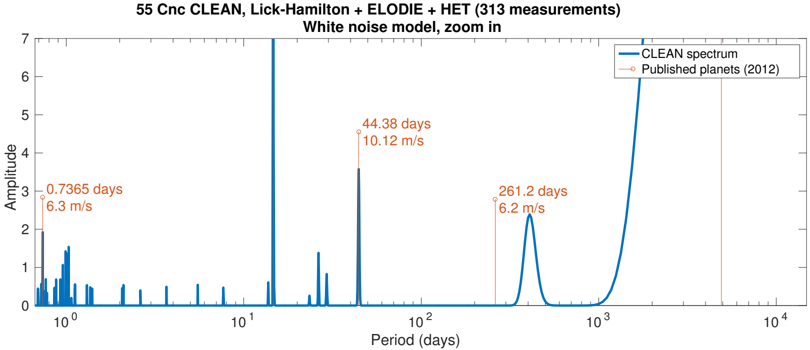


55 Cnc l1-periodogram, Lick-Hamilton + ELODIE + HET (313 measurements)
White noise model
Semi-Amplitude (m/s)
0
10
20
30
40
50
60
70
80
10^0
10^1
10^2
10^3
10^4
Period (days)

55 Cancri

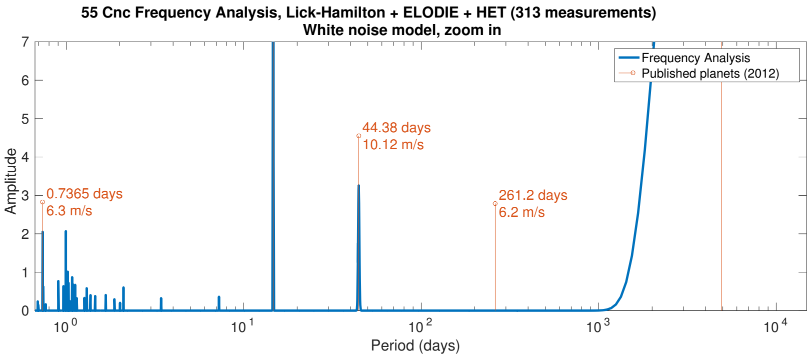

55 Cnc l1 periodogram, Lick-Hamilton + ELODIE + HET (313 measurements)
White noise model, zoom in
Instrumental offsets corrected
Published planets (2012)
Semi-Amplitude (m/s)
Period (days)
0.7365 days
6.3 m/s
0.99709 days
4 m/s
2.806 days
0.3 m/s
44.38 days
10.12 m/s
261.2 days
6.2 m/s
470 days
4.1 m/s
15
10
5
0
10^0
10^1
10^2
10^3
10^4

55 Cancri

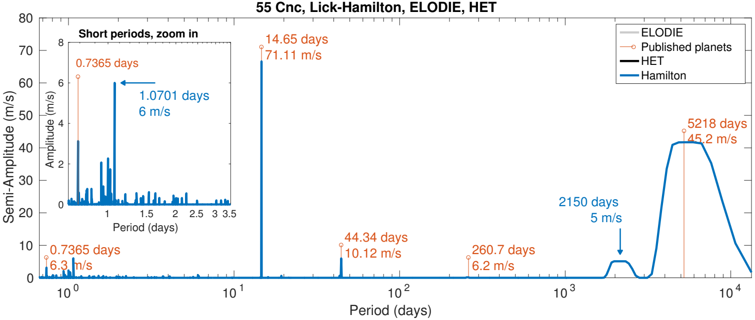


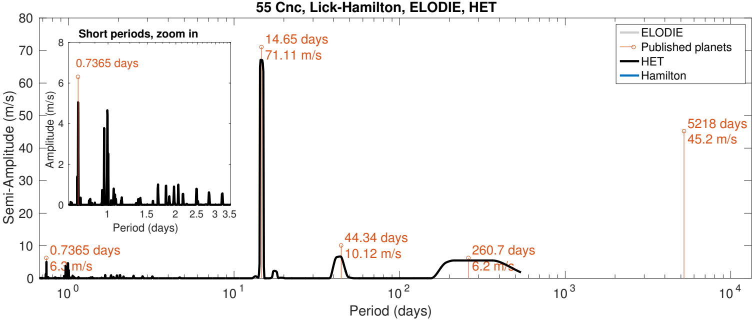
55 Cnc CLEAN, Lick-Hamilton + ELODIE + HET (313 measurements)
White noise model, zoom in
CLEAN spectrum
Published planets (2012)
0.7365 days
6.3 m/s
44.38 days
10.12 m/s
261.2 days
6.2 m/s
Amplitude
Period (days)
0
1
2
3
4
5
6
7
10^0
10^1
10^2
10^3
10^4

55 Cancri



55 Cnc Frequency Analysis, Lick-Hamilton + ELODIE + HET (313 measurements)
White noise model, zoom in
Frequency Analysis
Published planets (2012)
0.7365 days
6.3 m/s
44.38 days
10.12 m/s
261.2 days
6.2 m/s
Amplitude
Period (days)
0
1
2
3
4
5
6
7
10^0
10^1
10^2
10^3
10^4





Hara, Boué, Laskar, Correia 2016

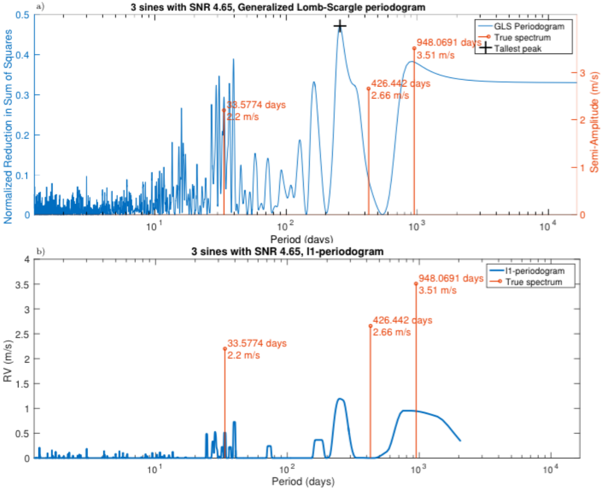
Examples of first wrong peak

Experiment

- 74 measurement dates of HD69830 spanning over 800 days
- Generate a sum of three sines with HD69830b, c, d amplitudes (3.51, 2.66, 2.2 m/s)
- 60 cm/s noise standard deviation (more precise than HARPS)
- Periods chosen randomly in $\log P$ between 1.2 and 2000 days

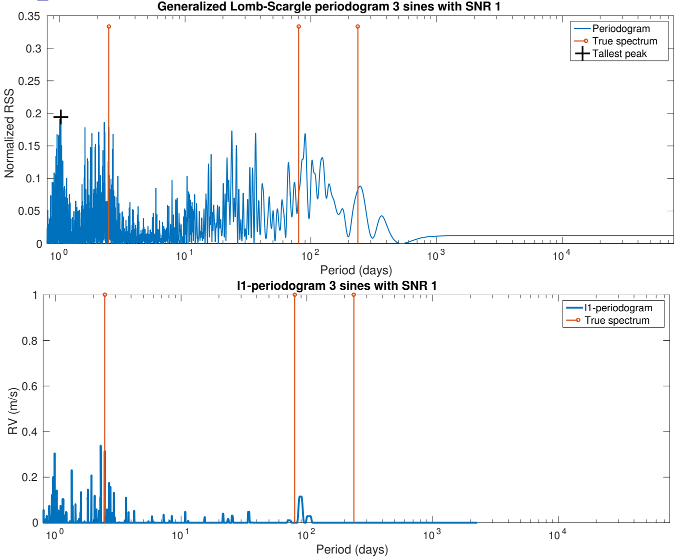
On 500 simulated systems : the Generalized Lomb-Scargle periodogram fails 33 times, and the ℓ_1-periodogram only twice

ℓ_1-periodogram failure



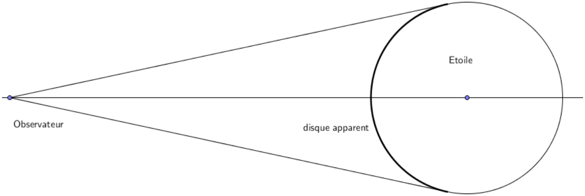
Hara, Boué, Laskar, Correia 2016

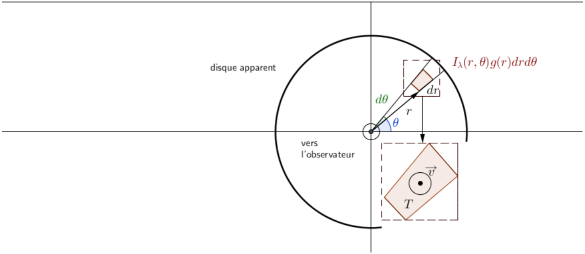
ℓ_1-periodogram failure

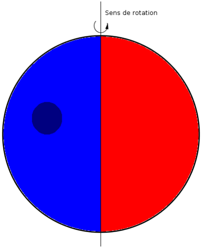


Hara, Boué, Laskar, Correia 2016

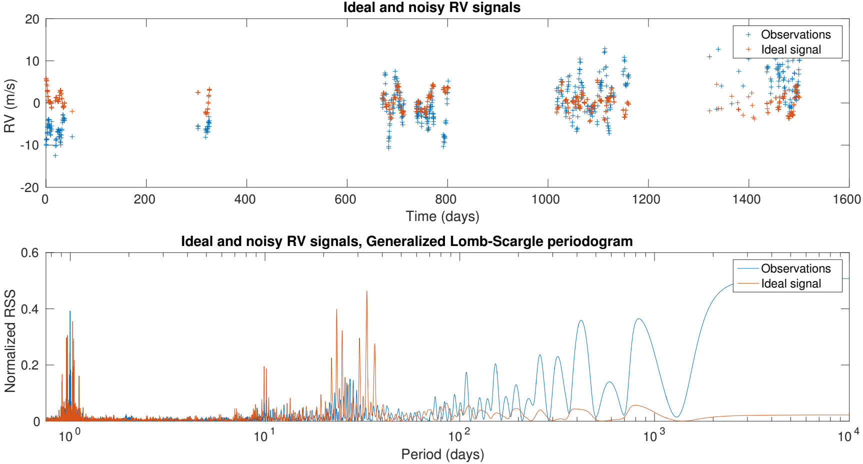
Stellar noise





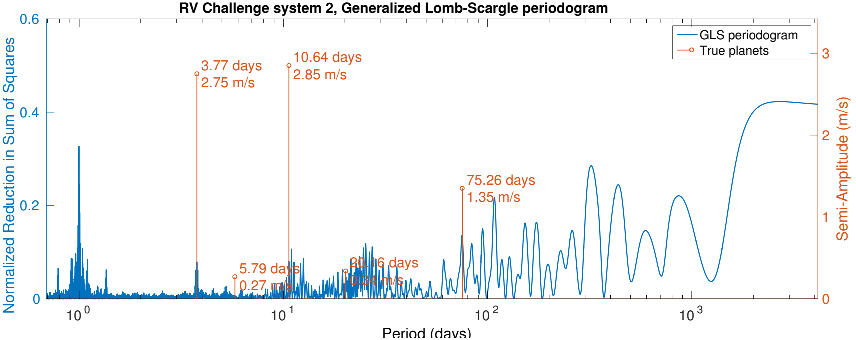


System with activity (simulated)

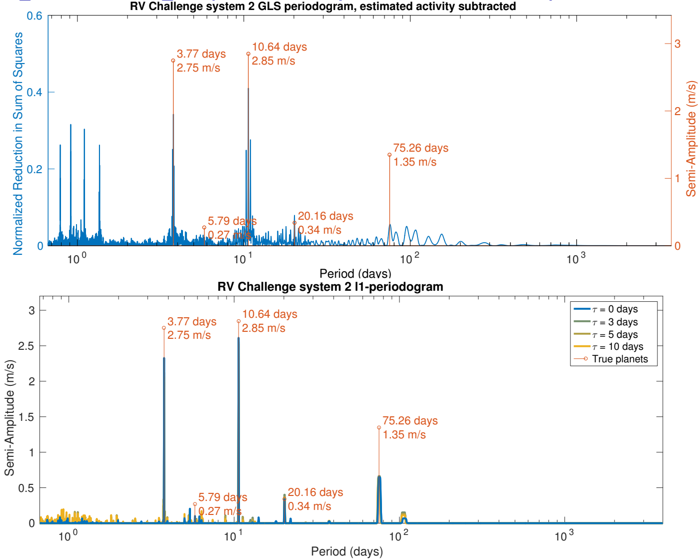


Issu du *Radial Velocity Fitting Challenge*, Dumusque et al 2016

RV Fitting Challenge



RV Fitting Challenge (Dumusque et al 2016)



No team found the 20.16 days periodicity during the Challenge

Outlook

Sparsity-based methods used for MP3, AAC, JPEG2000, electronics (e.g. Mishali et al 2008), image processing (e.g. Starck et al 2005)...
If you have a problem of de-noising or an inverse problem that can be recast as

$$y = Ax + \epsilon \text{ , a few entries of x are non-zero} \quad (7)$$

see An Introduction to Compressive Sampling (Candès 2008). Also for theory : Candès, Romberg Tao 2006 ; Donoho 2006

- Avoids fitting features one by one
- Faster than random searches, though less precise

Compressed sensing : infinite dictionary

Atomic norms

- Here we search a small number of sine, the dictionary is infinite. Chandrasekharan et al (2010), Candès et Fernandez-Granda (2012) et Tang, Bhaskar, Shah, Recht (2013) suggest to use the atomic norm :

$$\|y\|_{\mathcal{A}} = \inf \left\{ \sum_j |c_j|, y = \sum_j c_j a(\omega_j) \right\} \tag{8}$$

- When the observations are noisy (Candès et Fernandez-Granda 2013) :

$$u^\star = \underset{u \in \mathbb{C}^m}{\arg\min} \quad \|u - y\|_{\ell_2}^2 + \lambda \|u\|_{\mathcal{A}} \tag{AD$_\lambda$}$$

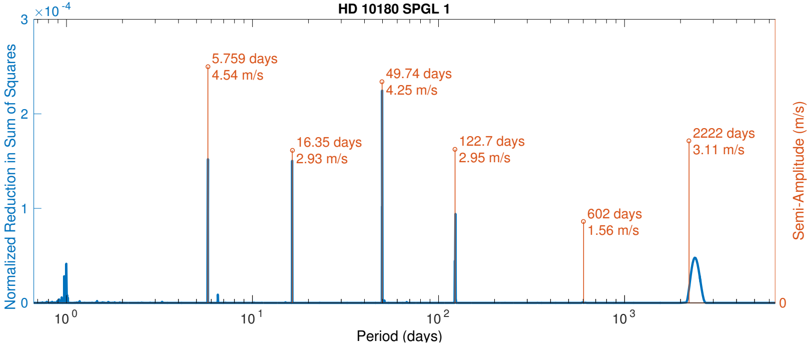
Compressed sensing : infinite dictionary

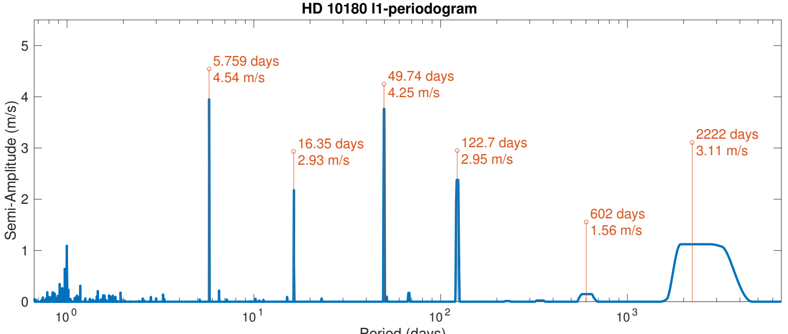
Implementation

- Several approaches : Matrix completion or discretization
- We choose discretization : we select vectors $(\mathrm{e}^{\mathrm{i}\omega\mathrm{t}})_{\omega==\omega_1..\omega_n}$ of dictionary $\mathcal{A}$

$$x^\star = \underset{x\in\mathbb{R}^n}{\arg\min} \quad \|x\|_{\ell_1} \quad \text{s. t.} \quad \|W(Ax-y)\|_{\ell_2} \leqslant \epsilon \qquad (\mathrm{BP}_{\epsilon,W})$$

  then "sliding mean" on $x^\star$
- Algorithms tested to solve $\mathrm{BP}_{\epsilon,W}$: ℓ_1-magic (Candès, Romberg, Tao 2006), SparseLab (Donoho et al 2006), NESTA (Becker, Bobin, Candès 2011) , CVX (Grant & Boyd 2008), Spectral Compressive Sampling (Duarte & Baraniuk 2013) et SPGL1 (van den Berg & Friedlander 2008)


HD 10180 SPGL 1
×10^-4
Normalized Reduction in Sum of Squares
Semi-Amplitude (m/s)
5.759 days
4.54 m/s
49.74 days
4.25 m/s
16.35 days
2.93 m/s
122.7 days
2.95 m/s
2222 days
3.11 m/s
602 days
1.56 m/s
Period (days)


HD 10180 l1-periodogram
Semi-Amplitude (m/s)
5.759 days
4.54 m/s
49.74 days
4.25 m/s
16.35 days
2.93 m/s
122.7 days
2.95 m/s
2222 days
3.11 m/s
602 days
1.56 m/s
Period (days)

Tolerance choice

$$x^\star = \underset{x \in \mathbb{R}^n}{\arg\min} \quad \|x\|_{\ell_1} \quad \text{s. t.} \quad \|W(Ax - y)\|_{\ell_2} \leqslant \epsilon \qquad (\mathrm{BP}_{\epsilon,W})$$

$$\|W(Ax - y)\|_{\ell_2}^2 = (y - Ax)^T V^{-1} (y - Ax) \qquad (9)$$

ϵ is chosen so that non noisy signal y_t verifies $\|W(Ax - y_t)\|_{\ell_2} \leqslant \epsilon$ with probability $1 - \alpha$:

$$F_{\chi^2_m}(\epsilon^2_{noise}) = 1 - \alpha \qquad (10)$$

Key for success

$$x^\star = \underset{x\in\mathbb{R}^n}{\arg\min} \quad \|x\|_{\ell_1} \quad \text{s. t.} \quad \|W(Ax - y(t))\|_{\ell_2} \leqslant \epsilon \qquad (\mathrm{BP}_{\epsilon,W})$$

Mutual coherence : maximum of correlation between two columns

$$\mu = \max_{k\neq l} |\langle A_k, A_l\rangle| \tag{11}$$

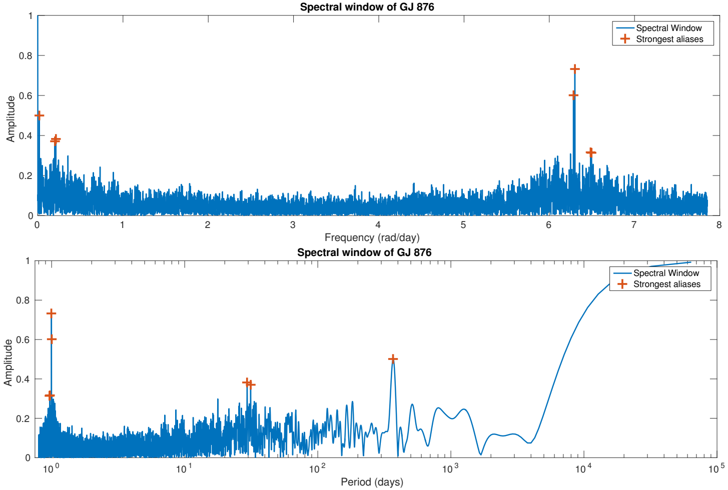
Theory garanties success for μ sufficiently small (Donoho et al 2006)
Since the k^{th} column $A_k = \mathrm{e}^{\mathrm{i}\omega_K t} / \sqrt{m}$

$$|\langle A_k A_l\rangle| = \frac{1}{m}\left|\sum_{i=1}^{m} \mathrm{e}^{\mathrm{i}(\omega_k - \omega_l)t_i}\right| = S_w(\omega_k - \omega_l) \tag{12}$$

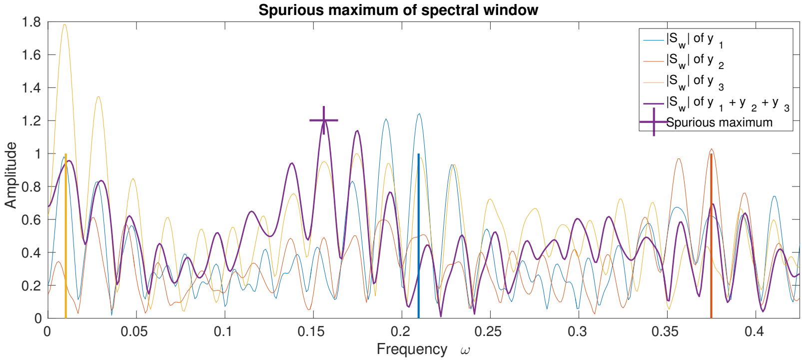
Which is the spectral window.
If $S_w(\Delta\omega) \gtrsim 0.7$ then it is hard to distinguish components at frequency ω and $\omega + \Delta\omega$

Spectral window

$$S_w(\Delta\omega) = \frac{1}{m}\left|\sum_{i=1}^{m} \mathrm{e}^{\mathrm{i}\Delta\omega t_i}\right| \tag{13}$$



Spectral window copies



Grid choice

Matrix A is made of $e^{-i\omega_k t}$ et $e^{i\omega_k t}$. How to choose ω_k ?
We take equispaced frequencies, then the following must be set

- Frequency spacing
- Minimum and maximum frequency

We make sure there exists a solution with the correct ℓ_0 norm :

$$\Delta\omega \leqslant \frac{4}{T} \arcsin \frac{\epsilon}{2\sqrt{\sum_{j=1}^{p} |c_j|^2}\sqrt{\sum_{k=1}^{m} \frac{1}{\sigma_k^2}}}. \tag{14}$$

$\omega_{min} = 0$ et $\omega_{max} = 1.5$ cycles/day or 0.95 cycles/day.

Weight matrix

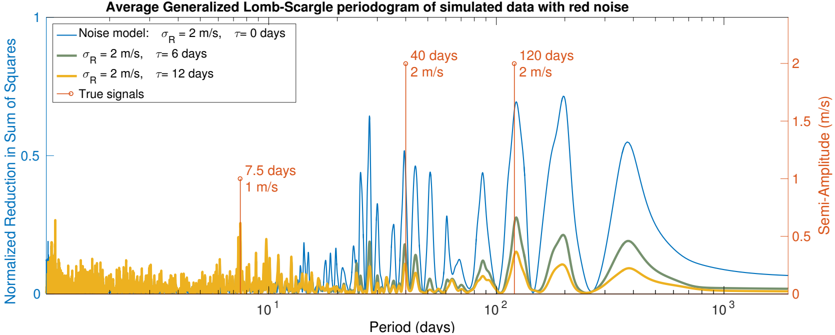
$$x^\star = \underset{x\in\mathbb{R}^n}{\arg\min} \quad \|x\|_{\ell_1} \quad \text{s. t.} \quad \|\mathbf{W}(Ax-y)\|_{\ell_2} \leqslant \epsilon \qquad (\mathrm{BP}_{\epsilon,W})$$

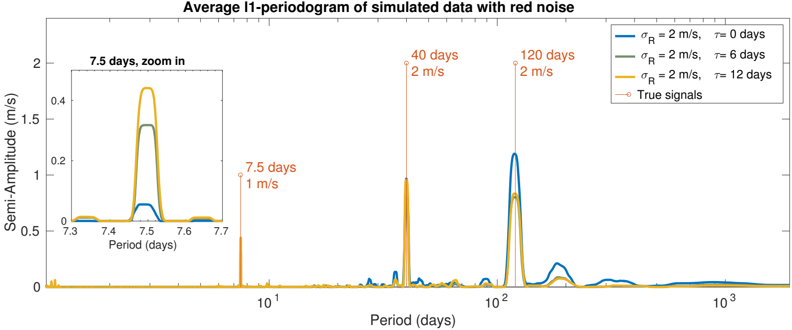
$W = V^{-\frac{1}{2}}$ where $V = \mathbb{E}\{\epsilon\epsilon^*\}$ is the covariance matrix.

Choix de W

A stationary noise is characterised by an autocorrelation function $R(|t-t'|) = \mathbb{E}\{\epsilon_t\epsilon_{t'}\}$. We have chosen :

$$\begin{aligned} R(\Delta t) &= \sigma_R^2\, \mathrm{e}^{-\frac{|\Delta t|}{\tau}}, \quad \Delta t \neq 0 \\ R(0) &= \sigma_W^2 + \sigma_R^2 \end{aligned} \qquad (15)$$

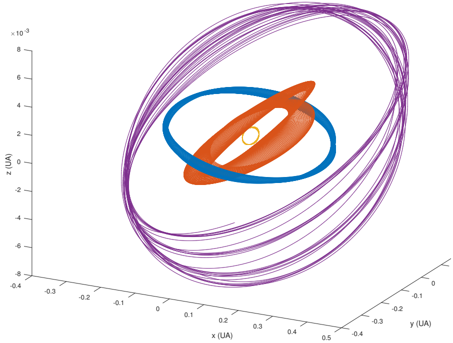

Average Generalized Lomb-Scargle periodogram of simulated data with red noise
Noise model: σ_R = 2 m/s, τ= 0 days
σ_R = 2 m/s, τ= 6 days
σ_R = 2 m/s, τ= 12 days
True signals
7.5 days
1 m/s
40 days
2 m/s
120 days
2 m/s
Normalized Reduction in Sum of Squares
Semi-Amplitude (m/s)
Period (days)


Average l1-periodogram of simulated data with red noise
σ_R = 2 m/s, τ= 0 days
σ_R = 2 m/s, τ= 6 days
σ_R = 2 m/s, τ= 12 days
True signals
7.5 days, zoom in
7.5 days
1 m/s
40 days
2 m/s
120 days
2 m/s
Semi-Amplitude (m/s)
Period (days)

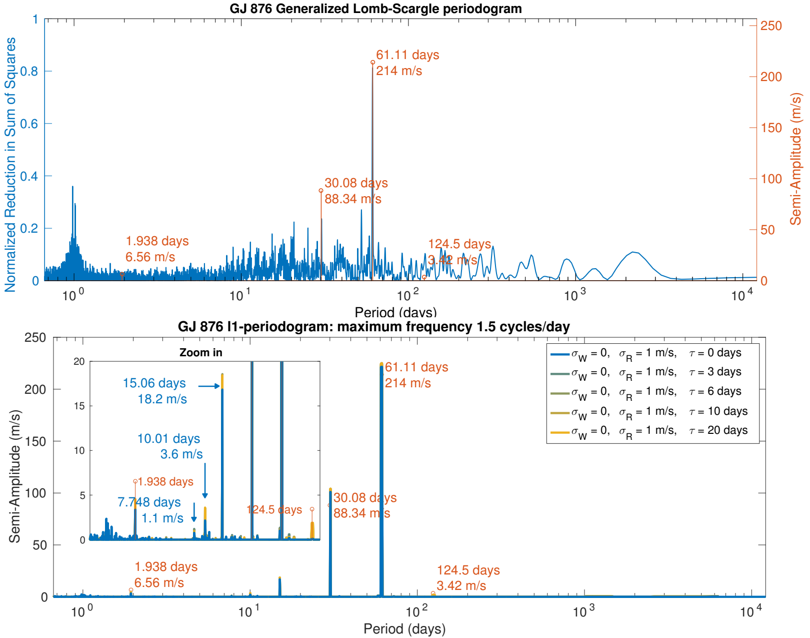
GJ 876

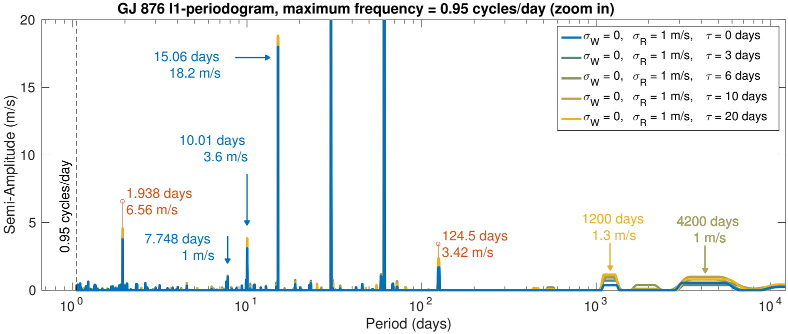
System with two planets in 2 :1 resonance

- Two Jupiters at 30 and 61 days are announced by Marcy et al 1998 and Delfosse et al 1998
- Detection of a Neptune at 1.94 days is reported in Rivera et al 2005.
- Uranus mass planet at 124 days reported (Rivera et al 2010)
- Correia et al 2010, Baluev 2011, Nelson et al 2016 fit Newtonian models

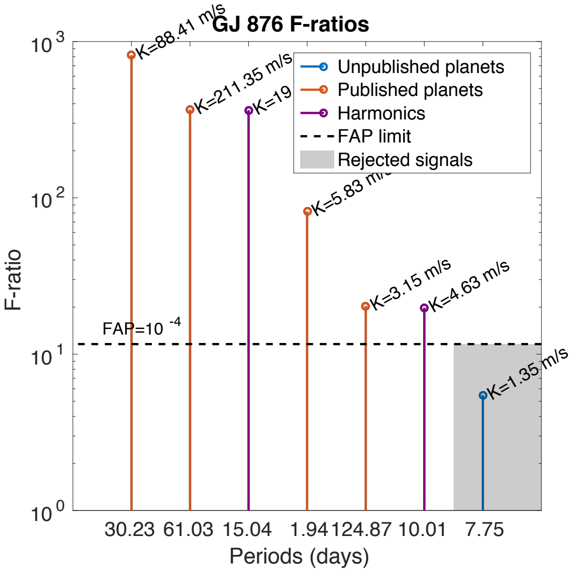


Simulated orbits on 100 years based on fitted parameters

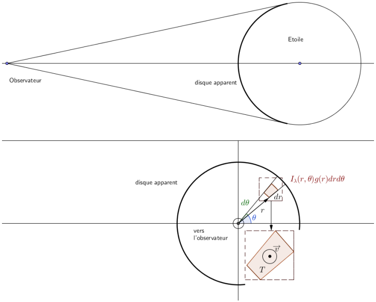

GJ 876 Generalized Lomb-Scargle periodogram
Normalized Reduction in Sum of Squares
Semi-Amplitude (m/s)
61.11 days
214 m/s
30.08 days
88.34 m/s
1.938 days
6.56 m/s
124.5 days
3.42 m/s
Period (days)
GJ 876 l1-periodogram: maximum frequency 1.5 cycles/day
Zoom in
15.06 days
18.2 m/s
10.01 days
3.6 m/s
7.748 days
1.1 m/s
$\sigma_W = 0$, $\sigma_R = 1$ m/s, $\tau = 0$ days
$\sigma_W = 0$, $\sigma_R = 1$ m/s, $\tau = 3$ days
$\sigma_W = 0$, $\sigma_R = 1$ m/s, $\tau = 6$ days
$\sigma_W = 0$, $\sigma_R = 1$ m/s, $\tau = 10$ days
$\sigma_W = 0$, $\sigma_R = 1$ m/s, $\tau = 20$ days

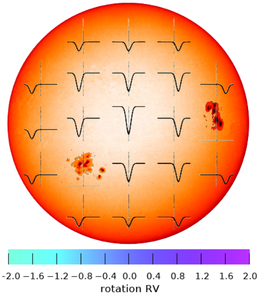


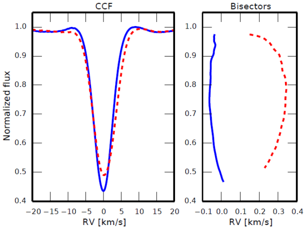
Hara, Boué, Laskar, Correia 2016

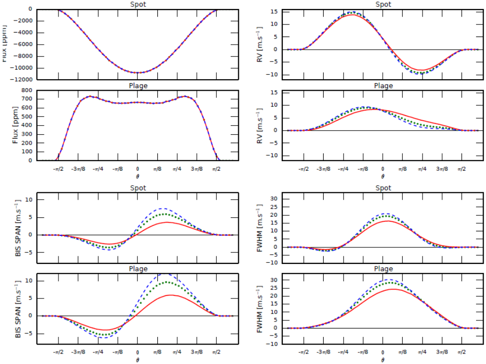


Hara, Boué, Laskar, Correia 2016









(Dumusque, Boisse, Santos 2014)

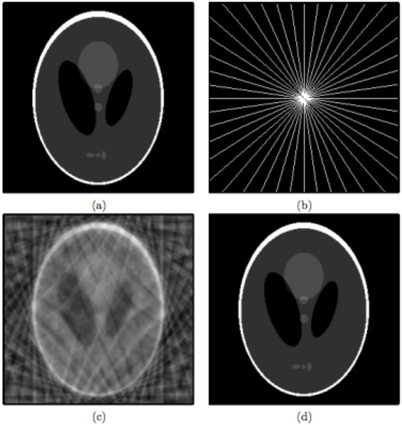
Conclusion for RV data analysis

Results

- A method easily implemented thanks to previous work
- Our contribution : adapted to correlated noise, automated tunings and smoothing (on-going theoretical work)
- Avoids spurious peaks in many situations, obvious when it fails
- Gives similar results to MCMC and Gaussian processes on RV Fitting Challenge (Dumusque 2016) but faster (30 secondes - 10 minutes) and without red noise assumptions.

How to use ℓ_1 ?

- Measures the system's difficulty, well suited for deciding if a target is worth further observations
- Possible short-cut to the solution
- Use statistical hypothesis testing to conclude (e.g. MCMC)



(a) (b) (c) (d)

Figure 1: Example of a simple recovery problem. (a) The Logan-Shepp phantom test image. (b) Sampling domain Ω in the frequency plane; Fourier coefficients are sampled along 22 approximately radial lines. (c) Minimum energy reconstruction obtained by setting unobserved Fourier coefficients to zero. (d) Reconstruction obtained by minimizing the total variation, as in (1.1). The reconstruction is an exact replica of the image in (a).

(Candès Romberg Tao 2005)

Thank you for your attention